

- 36 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: (OTHER THAN US) DARATECH PTY LTD and PIG RESEARCH
(US ONLY): MICHAEL PANACCIO and DETLEF HASSE

(ii) TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC
COMPOSITIONS

(iii) NUMBER OF SEQUENCES: 23

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: DAVIES COLLISON CAVE

(B) STREET: 1 LITTLE COLLINS STREET

(C) CITY: MELBOURNE

(D) STATE: VICTORIA

(E) COUNTRY: AUSTRALIA

(F) ZIP: 3000

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk

(B) COMPUTER: IBM PC compatible

(C) OPERATING SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) PCT INTERNATIONAL

(B) FILING DATE: 29-NOV-1996

(vii) PRIOR APPLICATION DATA:

09072574-092498

- 37 -

(A) APPLICATION NUMBER: PN6911/95

(B) FILING DATE: 30-NOV-1995

(A) APPLICATION NUMBER: PN6910/95

(B) FILING DATE: 30-NOV-1995

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: HUGHES DR, E JOHN L

(C) REFERENCE/DOCKET NUMBER: EJH/AF

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: +61 3 9254 2777

(B) TELEFAX: +61 3 9254 2770

090754-092498
864260-4252060

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1647 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1647

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	GCT	TCT	AAA	GAA	ATC	CTT	TTT	GAT	GCT	AAA	GCC	CGT	GAA	AAA	CTT	48
Met	Ala	Ser	Lys	Glu	Ile	Leu	Phe	Asp	Ala	Lys	Ala	Arg	Glu	Lys	Leu	
1				5					10					15		
TCA	CGA	GGT	GTA	GAT	AAA	CTT	GCA	AAT	GCT	GTT	AAA	GTA	ACA	CTT	GGA	96
Ser	Arg	Gly	Val	Asp	Lys	Leu	Ala	Asn	Ala	Val	Lys	Val	Thr	Leu	Gly	
			20					25					30			
CCT	AAA	GGC	CGT	AAT	GTC	GTT	ATT	GAA	AAG	TCT	TTT	GGT	TCC	CCA	GTT	144
Pro	Lys	Gly	Arg	Asn	Val	Val	Ile	Glu	Lys	Ser	Phe	Gly	Ser	Pro	Val	
		35					40					45				
ATT	ACA	AAA	GAT	GGT	GTA	TCT	GTT	GCA	AAA	GAA	ATT	GAA	CTT	GAA	GAT	192
Ile	Thr	Lys	Asp	Gly	Val	Ser	Val	Ala	Lys	Glu	Ile	Glu	Leu	Glu	Asp	
	50						55				60					
AAG	TTT	GAA	AAT	ATG	GGC	GCT	CAA	ATG	GTT	AAA	GAA	GTA	GCT	CCC	AAA	240
Lys	Phe	Glu	Asn	Met	Gly	Ala	Gln	Met	Val	Lys	Glu	Val	Ala	Pro	Lys	
65				70					75					80		

- 39 -

ACT AGC GAT ATT GCT GGT GAT GGA ACT ACA ACA GCA ACA GTC CTT GCA	288
Thr Ser Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala	
85 90 95	
CAA GCT ATT TAT CGT GAA GGT GTA AAA CTT GTA GCA GCT GGT CGT AAT	336
Gln Ala Ile Tyr Arg Glu Gly Val Lys Leu Val Ala Ala Gly Arg Asn	
100 105 110	
CCT ATG GCC ATT AAA CGT GGC ATA GAT AAA GCT GTT GTT GCT GTT ACT	38
Pro Met Ala Ile Lys Arg Gly Ile Asp Lys Ala Val Val Ala Val Thr	
115 120 125	
AAA GAA CTA AGC GAC ATT ACA AAG CCT ACT CGT GAC CAA AAA GAA ATA	432
Lys Glu Leu Ser Asp Ile Thr Lys Pro Thr Arg Asp Gln Lys Glu Ile	
130 135 140	
GCT CAA GTT GGA ACC ATT TCT GCA AAC TCT GAT ACA ACA ATA GGT AAT	480
Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Thr Thr Ile Gly Asn	
145 150 155 160	
ATC ATA GCT GAA GCT ATG GCT AAA GTT GGA AAA GGA GGT GTT ATC ACA	528
Ile Ile Ala Glu Ala Met Ala Lys Val Gly Lys Gly Gly Val Ile Thr	
165 170 175	
GTT GAG GAA GCT AAA GGT CTT GAA ACT ACA TTA GAT GTG GTT GAA GGA	576
Val Glu Glu Ala Lys Gly Leu Glu Thr Thr Leu Asp Val Val Glu Gly	
180 185 190	
ATG AAG TTT GAC CGT GGC TAC CTC TCT CCA TAC TTT GTA ACT AAT CCT	624
Met Lys Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Val Thr Asn Pro	
195 200 205	
GAG AAA ATG GTT TGT GAA CTT GAT AAC CCT TAT ATC CTT TGT AAT GAG	672
Glu Lys Met Val Cys Glu Leu Asp Asn Pro Tyr Ile Leu Cys Asn Glu	
210 215 220	
AAA AAG ATT ACT AGC ATG AAA GAC ATG CTA CCA ATC TTA GAA CAA GTT	720

09077574-092498

- 40 -

Lys Lys Ile Thr Ser Met Lys Asp Met Leu Pro Ile Leu Glu Gln Val
 225 230 235 240

 GCT AAA GTA AAC CGT CCA CTC CTT ATT ATT GCT GAA GAC GTA GAA GGT 768
 Ala Lys Val Asn Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
 245 250 255

 GAA GCA CTT GCA ACA CTT GTA GTC AAT AAG CTC CGT GGA GCA CTC CAA 816
 Glu Ala Leu Ala Thr Leu Val Val Asn Lys Leu Arg Gly Ala Leu Gln
 260 265 270

 GTT GTA GCC GTA AAA GCT CCT GGT TTT GGT GAA CGC CGT AAA GCT ATG 864
 Val Val Ala Val Lys Ala Pro Gly Phe Gly Glu Arg Arg Lys Ala Met
 275 280 285

 CTT GAA GAT ATT GCT ATC CTT ACT GGA GGA GAA GCA ATA TTT GAA GAT 912
 Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly Glu Ala Ile Phe Glu Asp
 290 295 300

 CGT GGT ATA AAG CTT GAA AAT GTA AGC TTG TCT TCT TTA GGA ACA GCT 960
 Arg Gly Ile Lys Leu Glu Asn Val Ser Leu Ser Ser Leu Gly Thr Ala
 305 310 315 320

 AAA CGT GTA GTT ATT GAC AAA GAA AAT ACT ACT ATC GTT GAT GGT GCT 1008
 Lys Arg Val Val Ile Asp Lys Glu Asn Thr Thr Ile Val Asp Gly Ala
 325 330 335

 GGA AAA TCA GAA GAT ATT AAA GCT CGA GTT AAA CAA ATT CGT GCA CAA 1056
 Gly Lys Ser Glu Asp Ile Lys Ala Arg Val Lys Gln Ile Arg Ala Gln
 340 345 350

 ATT GAA GAA ACA AGC TCA GAT TAT GAT CGT GAA AAA CTT CAA GAA CGT 1104
 Ile Glu Glu Thr Ser Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg
 355 360 365

 CTT GCA AAA CTT GTT GGT GGA GTA GCT GTT ATC CAT GTT GGA GCT GCT 1152
 Leu Ala Lys Leu Val Gly Gly Val Ala Val Ile His Val Gly Ala Ala

0907574-092498

- 41 -

370 375 380

ACT GAA ACT GAA ATG AAA GAG AAG AAG GAT CGT GTA GAA GAT GCT CTA 1200
 Thr Glu Thr Glu Met Lys Glu Lys Lys Asp Arg Val Glu Asp Ala Leu
 385 390 395 400

AAT GCA ACA AGA GCT GCG GTT GAA GAA GGT ATT GTC CCT GGT GGT GGT 1248
 Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly
 405 410 415

ACT GCT TTT GTC CGC TCC ATT AAA GTC CTT GAT GAT ATT AAA CCT GCT 1296
 Thr Ala Phe Val Arg Ser Ile Lys Val Leu Asp Asp Ile Lys Pro Ala
 420 425 430

GAT GAT GAT GAA CTT GCT GGA CTT AAT ATC ATC CGT CGT TCT CTT GAA 1344
 Asp Asp Asp Glu Leu Ala Gly Leu Asn Ile Ile Arg Arg Ser Leu Glu
 435 440 445

GAG CCT TTA CGT CAA ATT GCT GCA AAT GCT GGC TAT GAA GGT TCT ATT 1392
 Glu Pro Leu Arg Gln Ile Ala Ala Asn Ala Gly Tyr Glu Gly Ser Ile
 450 455 460

GTT GTA GAA AAA GTT CGT GAA CCA AAA GAT GGT TTT GGA TTT AAT GCT 1440
 Val Val Glu Lys Val Arg Glu Pro Lys Asp Gly Phe Gly Phe Asn Ala
 465 470 475 480

GCA TCA GGA GAA TAT GAA GAC CTT ATT AAA GCT GGT GTC ATT GAT CCT 1488
 Ala Ser Gly Glu Tyr Glu Asp Leu Ile Lys Ala Gly Val Ile Asp Pro
 485 490 495

AAA AAA GTT ACA CGT ATT GCA TTA CAA AAT GCA GCA TCA GTA GCC TCC 1536
 Lys Lys Val Thr Arg Ile Ala Leu Gln Asn Ala Ala Ser Val Ala Ser
 500 505 510

TTA CTT CTA ACT ACA GAA TGC GCT ATT GCT GAA AAA CCA GAA CCT AAA 1584
 Leu Leu Leu Thr Thr Glu Cys Ala Ile Ala Glu Lys Pro Glu Pro Lys
 515 520 525

864260 42542060

AAA GAT ATG CCT ATG CCT GGC GGT GGT ATG GGT GGT ATG GGT GGT ATG 1632
Lys Asp Met Pro Met Pro Gly Gly Gly Met Gly Gly Met Gly Gly Met
530 535 540

GAC GGT ATG TAC TAG 1647
Asp Gly Met Tyr
545

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 548 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Ser Lys Glu Ile Leu Phe Asp Ala Lys Ala Arg Glu Lys Leu
1 5 10 15

Ser Arg Gly Val Asp Lys Leu Ala Asn Ala Val Lys Val Thr Leu Gly
20 25 30

Pro Lys Gly Arg Asn Val Val Ile Glu Lys Ser Phe Gly Ser Pro Val
35 40 45

Ile Thr Lys Asp Gly Val Ser Val Ala Lys Glu Ile Glu Leu Glu Asp
50 55 60

Lys Phe Glu Asn Met Gly Ala Gln Met Val Lys Glu Val Ala Pro Lys
65 70 75 80

Thr Ser Asp Ile Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala
85 90 95

[illegible]

- 43 -

Gln Ala Ile Tyr Arg Glu Gly Val Lys Leu Val Ala Ala Gly Arg Asn
 100 105 110

Pro Met Ala Ile Lys Arg Gly Ile Asp Lys Ala Val Val Ala Val Thr
 115 120 125

Lys Glu Leu Ser Asp Ile Thr Lys Pro Thr Arg Asp Gln Lys Glu Ile
 130 135 140

Ala Gln Val Gly Thr Ile Ser Ala Asn Ser Asp Thr Thr Ile Gly Asn
 145 150 155 160

Ile Ile Ala Glu Ala Met Ala Lys Val Gly Lys Gly Gly Val Ile Thr
 165 170 175

Val Glu Glu Ala Lys Gly Leu Glu Thr Thr Leu Asp Val Val Glu Gly
 180 185 190

Met Lys Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Val Thr Asn Pro
 195 200 205

Glu Lys Met Val Cys Glu Leu Asp Asn Pro Tyr Ile Leu Cys Asn Glu
 210 215 220

Lys Lys Ile Thr Ser Met Lys Asp Met Leu Pro Ile Leu Glu Gln Val
 225 230 235 240

Ala Lys Val Asn Arg Pro Leu Leu Ile Ile Ala Glu Asp Val Glu Gly
 245 250 255

Glu Ala Leu Ala Thr Leu Val Val Asn Lys Leu Arg Gly Ala Leu Gln
 260 265 270

Val Val Ala Val Lys Ala Pro Gly Phe Gly Glu Arg Arg Lys Ala Met
 275 280 285

Leu Glu Asp Ile Ala Ile Leu Thr Gly Gly Glu Ala Ile Phe Glu Asp

090754-09498
 864260-42524060

- 44 -

290	295	300
Arg Gly Ile Lys Leu Glu Asn Val Ser Leu Ser Ser Leu Gly Thr Ala		
305	310	315 320
Lys Arg Val Val Ile Asp Lys Glu Asn Thr Thr Ile Val Asp Gly Ala		
325	330	335
Gly Lys Ser Glu Asp Ile Lys Ala Arg Val Lys Gln Ile Arg Ala Gln		
340	345	350
Ile Glu Glu Thr Ser Ser Asp Tyr Asp Arg Glu Lys Leu Gln Glu Arg		
355	360	365
Leu Ala Lys Leu Val Gly Gly Val Ala Val Ile His Val Gly Ala Ala		
370	375	380
Thr Glu Thr Glu Met Lys Glu Lys Lys Asp Arg Val Glu Asp Ala Leu		
385	390	395 400
Asn Ala Thr Arg Ala Ala Val Glu Glu Gly Ile Val Pro Gly Gly Gly		
405	410	415
Thr Ala Phe Val Arg Ser Ile Lys Val Leu Asp Asp Ile Lys Pro Ala		
420	425	430
Asp Asp Asp Glu Leu Ala Gly Leu Asn Ile Ile Arg Arg Ser Leu Glu		
435	440	445
Glu Pro Leu Arg Gln Ile Ala Ala Asn Ala Gly Tyr Glu Gly Ser Ile		
450	455	460
Val Val Glu Lys Val Arg Glu Pro Lys Asp Gly Phe Gly Phe Asn Ala		
465	470	475 480
Ala Ser Gly Glu Tyr Glu Asp Leu Ile Lys Ala Gly Val Ile Asp Pro		
485	490	495

- 45 -

Lys Lys Val Thr Arg Ile Ala Leu Gln Asn Ala Ala Ser Val Ala Ser
 500 505 510

Leu Leu Leu Thr Thr Glu Cys Ala Ile Ala Glu Lys Pro Glu Pro Lys
 515 520 525

Lys Asp Met Pro Met Pro Gly Gly Gly Met Gly Gly Met Gly Gly Met
 530 535 540

Asp Gly Met Tyr
 545

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..306

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG AAC CTG AAA CCT TTG AAT GAC CGT GTT TTA GTA AAA CGT CTT GAA 48
 Met Asn Leu Lys Pro Leu Asn Asp Arg Val Leu Val Lys Arg Leu Glu
 1 5 10 15

TCT GAA GAA AAA ACA GCT GGT GGA CTC TAT ATC CCT GAT ACT GCT AAA 96
 Ser Glu Glu Lys Thr Ala Gly Gly Leu Tyr Ile Pro Asp Thr Ala Lys
 20 25 30

86426042522060

- 46 -

GAA AAA CCA TCT CGT GGT GAA GTT GTT GCT GTT GGA CCT GGT AAA CAT 144
 Glu Lys Pro Ser Arg Gly Glu Val Val Ala Val Gly Pro Gly Lys His
 35 40 45

ACA GAT GAT GGT AAA TTA ATA CCT ATG GCT GTA AAA GCA GGA GAT ACA 192
 Thr Asp Asp Gly Lys Leu Ile Pro Met Ala Val Lys Ala Gly Asp Thr
 50 55 60

GTT CTT TTT AAT AAG TAT GCA GGA ACA GAA GTA AAG CTT GAT GGT GTA 240
 Val Leu Phe Asn Lys Tyr Ala Gly Thr Glu Val Lys Leu Asp Gly Val
 65 70 75 80

GAG CAT CTA GTT ATG CGT GAA GAT GAC ATC CTA GCT GTT ATT ACT GGA 288
 Glu His Leu Val Met Arg Glu Asp Asp Ile Leu Ala Val Ile Thr Gly
 85 90 95

GAA ACT GGC CGC AAG TGA 306
 Glu Thr Gly Arg Lys *
 100

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 101 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Asn Leu Lys Pro Leu Asn Asp Arg Val Leu Val Lys Arg Leu Glu
 1 5 10 15

Ser Glu Glu Lys Thr Ala Gly Gly Leu Tyr Ile Pro Asp Thr Ala Lys

0907574-092498

3.

3.

3.

3.

- 3.

3.

3.

AACTCCTGGT	CTATCAAGAT	CAACTAAAAA	ATATTCTTTA	TCTAATAGTT	50
GCTCAAAAAT	AATTGTACCT	ACAGGTAAAT	GAAGAATCAA	ATCTTCCCCT	100
TTTTTACCAT	GACGCTGGCT	CCCTTTACCA	CCTTCTCCAT	TTTGAGCTCT	150
ATAGTGACGT	TGCACACGAA	AATCATAAAG	GGTTAACAAA	CGTGAATCAG	200
CTTTAAAAAT	TATATTACCT	CCATCTCCTC	CATCCCCTCC	ATTAGGTCCA	250

- 48 -

CCTTTAGGTA	TAAACTTTTC	GCGTCTAAAT	GAAACACATC	CATTTCCACC	300
TTTTCTGCG	CTCACGCTAA	TAGTTACTTC	ATCAACAAAA	CGCATGATTA	350
TCCTTTCAAT	AACAAATATC	TATTCAATAC	TGTTACTAAC	TTGTTTACTG	400
TTTTTTCTAG	AAAATTACCT	GGCTAATTAT	TATAGTTATA	TCTAGATTAA	450
TGAAAAAGGA	AGAAGTCATT	ACACTCCTTC	CTTATTAATA	GAATCCTGGA	500
ATAATTATTA	TACGGTGGGT	TGTATATGCA	CTCTACTATA	TCTTTTACAT	550
TTACGAAAAT	ATGTTTCATA	AGTTACTATA	CCATTAACTT	TTGCAAATAA	600
AGTATAGTCT	CTTCCCATT	CAACATTTTC	TCCAGGATGA	ATTTTTGTAC	650
CTAGTTGACG	AACAAGGATA	TTGCCTGCCA	AGACTTTCTG	GCCGCCGAAA	700
CGCTTTATAC	CACGACGTTG	TCCTGGACTA	TCTCTACCAT	TGCGAGAACT	750
TCCACCAGCT	TTCTTATGGG	CCATTTTAAT	ATCTCCTTAA	AGCTGAATAC	800
CTGTTACTTT	TAGAGCTGTA	TAGTCTTGAC	GATGACCTTG	GAGTTTACGT	850
GAGTCATTTT	TTCTCCACTT	TTTAAAAACA	AGAATTTTTT	TATCACGACC	900
ATGCTCAAGA	ACTTTAGCTA	TAACTTTAGC	ATTATTAATA	TATGGTGTTT	950
CAATTTGAGG	AGATGAACCA	CCAATCATAA	AAATTTTATC	AAAAAAAATT	1000
TCTGTTCCAA	CTTCAGCGTC	TATTTTAGAA	ACAAAAATTT	TAGAACCCTC	1050
TTCAACACAG	AATTGTTTTT	CACCAGCTTC	AATAATTGCG	TACATAAATA	1100
ATGTGCCTCC	CAAAAAAGAC	AAGAAATACT	AATTTGATAT	TTTCAATATT	1150
GTCAAGTAGG	AACTTTATCT	TTAGAATGTT	AGATGTAACA	ATTTTTTTAG	1200
AAAAAAAATA	TTTTCAATAC	AATAGGAAAA	GAGGAAAAAA	AAAAAGATTT	1250
TTAGAAAAAA	TTTTTATTTT	TCCAAAAAAT	GCAAAAAATAT	AAAAAATTCT	1300
AATAGGATAG	AAGTTATTAC	TGTATTGATT	TTCAAGACTT	ACTTAAAAAT	1350
TTTTATAAAA	AAATTTGCAT	TCCCCTCTTC	CCAATTCCCA	TAGAGAAGAT	1400
TATTTATCCT	AACGATTGGT	GGACGCTAAG	TCCCTGCTGT	TTTGATTATA	1450
TATCAAATGT	TGAAACAAAT	TTTGTTTAGT	TTCTTTTTGT	ACTCTAAAAA	1500
GAAGACAAAA	AATTCTTTAT	AAACTGTACA	CTCTAAACAA	AATAGTTCAC	1550
AATAAACAGC	AATACATTAT	AATTAATTGG	AGGATACTAT	TGTCATGAAC	1600
CTGAAACCTT	TGAATGACCG	TGTTTTAGTA	AAACGTCTTG	AATCTGAAGA	1650
AAAAACAGCT	GGTGGACTCT	ATATCCCTGA	TACTGCTAAA	GAAAAACCAT	1700
CTCGTGGTGA	AGTTGTTGCT	GTTGGACCTG	GTAAACATAC	AGATGATGGT	1750
AAATTAATAC	CTATGGCTGT	AAAAGCAGGA	GATACAGTTC	TTTTTAATAA	1800
GTATGCAGGA	ACAGAAGTAA	AGCTTGATGG	TGTAGAGCAT	CTAGTTATGC	1850
GTGAAGATGA	CATCCTAGCT	GTTATTACTG	GAGAAACTGG	CCGCAAGTGA	1900
AAAAGGCGTA	AATAAAAAAG	TCGGTGATCT	TTAATAATTT	TATTCAGTTA	1950
TAATGAAAAC	ACTAATTACA	CGCACTCTCT	GAGAATTTTC	TCAGAAAACT	2000
ATATTTAACA	ATTCTAAAAT	CGATATGTTT	TTAGGAGGAA	AACCCTAATG	2050
GCTTCTAAAG	AAATCCTTTT	TGATGCTAAA	GCCCGTGAAA	AACTTTCACG	2100

09072574-09498

AGGTGTAGAT AACTTGCAA ATGCTGTAA AGTAACACTT GGACCTAAAG 2150
GCCGTAATGT CGTTATTGAA AAGTCTTTTG GTTCCCCAGT TATTACAAA 2200
GATGGTGTAT CTGTTGCAA AGAAATTGAA CTTGAAGATA AGTTTGAAA 2250
TATGGGCGCT CAAATGGTTA AAGAAGTAGC TCCCAAACT AGCGATATTG 2300
CTGGTGATGG AACTACAACA GCAACAGTCC TTGCACAAGC TATTTATCGT 2350
GAAGGTGTAA AACTTGTAGC AGCTGGTCGT AATCCTATGG CCATTAAACG 2400
TGGCATAGAT AAAGCTGTTG TTGCTGTTAC TAAAGAACTA AGCGACATTA 2450
CAAAGCCTAC TCGTGACCAA AAAGAAATAG CTCAAGTTGG AACCATTCT 2500
GCAAACTCTG ATACAACAAT AGGTAATATC ATAGCTGAAG CTATGGCTAA 2550
AGTTGGAAAA GGAGGTGTTA TCACAGTTGA GGAAGCTAAA GGTCTTGAAA 2600
CTACATTAGA TGTGGTTGAA GGAATGAAGT TTGACCGTGG CTACCTCTCT 2650
CCATACTTTG TAACTAATCC TGAGAAAATG GTTGTGAAC TTGATAACCC 2700
TTATATCCTT TGTAATGAGA AAAAGATTAC TAGCATGAAA GACATGCTAC 2750
CAATCTTAGA ACAAGTTGCT AAAGTAAACC GTCCACTCCT TATTATTGCT 2800
GAAGACGTAG AAGGTGAAGC ACTTGCAACA CTTGTAGTCA ATAAGCTCCG 2850
TGGAGCACTC CAAGTTGTAG CCGTAAAAGC TCCTGGTTTT GGTGAACGCC 2900
GTAAAGCTAT GCTTGAAGAT ATTGCTATCC TTAAGCTTGT CTTCTTTAGG 2950
TTTGAAGATC GTGGTATAAA GCTTGAAAAT GTAAGCTTGT CTTCTTTAGG 3000
AACAGCTAAA CGTGTAGTTA TTGACAAAGA AAATACTACT ATCGTTGATG 3050
GTGCTGGAAA ATCAGAAGAT ATTAAAGCTC GAGTTAAACA AATTCGTGCA 3100
CAAATTGAAG AAACAAGCTC AGATTATGAT CGTGAAAAC TTCAAGAACG 3150
TCTTGCAAAA CTTGTTGGTG GAGTAGCTGT TATCCATGTT GGAGCTGCTA 3200
CTGAAACTGA AATGAAAGAG AAGAAGGATC GTGTAGAAGA TGCTCTAAAT 3250
GCAACAAGAG CTGCGGTTGA AGAAGGTATT GTCCCTGGTG GTGGTACTGC 3300
TTTTGTCCGC TCCATTAAAG TCCTTGATGA TATTAAACCT GCTGATGATG 3350
ATGAACTTGC TGGACTTAAT ATCATCCGTC GTTCTCTTGA AGAGCCTTTA 3400
CGTCAAATTG CTGCAAATGC TGGCTATGAA GGTCTATTG TTGTAGAAAA 3450
AGTTCGTGAA CCAAAGATG GTTTTGGATT TAATGCTGCA TCAGGAGAAT 3500
ATGAAGACCT TATTAAAGCT GGTGTCATTG ATCCTAAAA AGTTACACGT 3550
ATTGCATTAC AAAATGCAGC ATCAGTAGCC TCCTTACTTC TAACTACAGA 3600
ATGCGCTATT GCTGAAAAAC CAGAACCTAA AAAAGATATG CCTATGCCTG 3650
GCGGTGGTAT GGGTGGTATG GGTGGTATGG ACGGTATGTA CTAGTCCTAT 3700
CTTCAGTACA ACTTAGATGT ATAAAAACCC CAGAAGCAAT GCTTCCGGGG 3750
TTTTATACTT TCAGCATAAA AAATTAATAT TTAATATACA GACACATTAT 3800
TTTGGTATTT ATTATTTATT ATGATCAAAT ATATAGACTG GATACAAAA 3850
ACAACAATGA TGTTTAAAAA GGCAGGGATA GATTCACCAA AACTCTCTGC 3900
AGAACTTATA TTAAGTCATG TTTTAAATAT TACACGATTA CAAATAATA 3950

964260-422260

WO 97/20050

- 50 -

TGACTCCTTT TGAACCTATT CCAACTAATA GCTACTCAAC GCTTAATGAT 4000
 ATCATGTTAA GAAGACTCCA TGGAGAACCA ATTGCATATC TCACAGGGAA 4050
 AAAAGAATTT TTTTCACGAG AATTAAAGT CACTCAAGCC ACACCTATCC 4100
 CTCGCCCAGA GACAGAGTTA CTTATAGAAT TTGTATTAAA CCATATTAAC 4150
 CCAACACAAC AAATATACTT TGCAGACTTA GGTACAGGTA GTGGGTGTAT 4200
 TGCAATTACA CTAGCTGCTG AAAGAAAAA TTGGTTAGGT ATTGCTACTG 4250
 ATATCTCTAG TGAAGCATTAA AAAATAGCTA AACTTAATAG TTTAAAAAAT 4300
 AACACTCATA GTCAACTACA GTTCTTCAA TCAGATTTTA CACAACCACT 4350
 CTGTCTACCC TCTTCATTAG ACTTATATAT CAGTAATCCT CCATATATAA 4400
 GTGAAAATGA ACTGACCTCT CTTCCGCATG AAGTAATATC TTTTGAACCT 4450
 AAAATAGCTC TTACACCACA TAAATGTATT CATCTTGATG AAATAAATAC 4500
 CGTTTTACAC TGCTATAAAA AAATTATTAC CCAAGCAGAG ATATCCCTTA 4550
 AGCCTGGAGG AATAATAATT TTAGAACATG GAGCAACACA AGCAGAAGCT 4600
 ATCTTATTGT TGTTAAAAA CAACATATGG ACAAATGTAA TAAGTCATAC 4650
 TGATCTTACA AATAAAAATC GTTTTATTAC AGCATATAAG TATAAAATAT 4700
 AACTTAATTA TGTTGkagAa AAAACAAAA ATAAAAATAA GATATtAAaT 4750
 ATTTtttttA aTAAAATTAA GCAAtTACTA ATATCTTTTT TTGGrTCGtt 4800
 yaTtGsATwA GAACTTTGG rGGrTrrCTa TGAACAAACA ACCATnCAAC 4850
 GGCCAAATAC ATnnCAGGnT TGGGGTCATA GGGGCCACGC TTTATGTACG 4900
 TACAACCCCh ACTGAAATTC TGGnTTGnTT TGGGGGnAA nTGGGTATCG 4950
 CAACnCTnTC CCCCCCCCCT GG 4972

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 569 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 209..569

- 51 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGTTAAAAAG TAAGGAGAAA AGGTTGGTTA AACCAAGTTT AAAAAATTAA TTTTTTTT	60
TTACCCAAAA AAGTTTATTA GATTAAGTAA TATTAATTTG GCCCAAAAAT TTTTTTGGGC	120
ATGGGTTTTT TGCTTTTAAA ATAGAGATGT GTAGGTAACA TTTTTCCTC CATGAAATTA	180
TTTTTTAGGA GATGTTATCA TGATGGGG AGT TTG TTT ATT GNT GCG AAC AGG	232
Ser Leu Phe Ile Xaa Ala Asn Arg	
1 5	
TAT GAA AAC CCA TAG NAC AGG GNT GGT ACT GTC TCC AAT AAT ATT GCT	280
Tyr Glu Asn Pro * Xaa Arg Xaa Gly Thr Val Ser Asn Asn Ile Ala	
10 15 20	
AAC GCA AAT ACC ATT GGG TAT AAG CAG CAA CAG GTA GTG TTT CAA GAC	328
Asn Ala Asn Thr Ile Gly Tyr Lys Gln Gln Gln Val Val Phe Gln Asp	
25 30 35 40	
CTG TTT AGT CAA GAT TTA GCA ATA GGT TTT ACT GGA AGT CAG GGG CCA	376
Leu Phe Ser Gln Asp Leu Ala Ile Gly Phe Thr Gly Ser Gln Gly Pro	
45 50 55	
AAC CAG GCT GGT ATG GGA GCA CAG GTG GGA AGT GTT CGC ACA ATT TTT	424
Asn Gln Ala Gly Met Gly Ala Gln Val Gly Ser Val Arg Thr Ile Phe	
60 65 70	
ACA CAG GGT GCT TTT GAA CCT GGC AAT AGT GTA ACA GAT CCT GCT ATT	472
Thr Gln Gly Ala Phe Glu Pro Gly Asn Ser Val Thr Asp Pro Ala Ile	
75 80 85	
GGT GGA AAA GGT TTT TTT CAG GTT ACA TTA GAG GAT AAA GTA CAC TAT	520
Gly Gly Lys Gly Phe Phe Gln Val Thr Leu Glu Asp Lys Val His Tyr	
90 95 100	

09077574-092498

WO 97/20050

- 52 -

ACA CGA GCA GGG AAT TTT CGT TTT ACT CAA GAT GGT TTT TTA AAT GAT C 569
 Thr Arg Ala Gly Asn Phe Arg Phe Thr Gln Asp Gly Phe Leu Asn Asp
 105 110 115 120

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Ser Leu Phe Ile Xaa Ala Asn Arg Tyr Glu Asn Pro * Xaa Arg Xaa
 1 5 10 15
 Gly Thr Val Ser Asn Asn Ile Ala Asn Ala Asn Thr Ile Gly Tyr Lys
 20 25 30
 Gln Gln Gln Val Val Phe Gln Asp Leu Phe Ser Gln Asp Leu Ala Ile
 35 40 45
 Gly Phe Thr Gly Ser Gln Gly Pro Asn Gln Ala Gly Met Gly Ala Gln
 50 55 60
 Val Gly Ser Val Arg Thr Ile Phe Thr Gln Gly Ala Phe Glu Pro Gly
 65 70 75 80
 Asn Ser Val Thr Asp Pro Ala Ile Gly Gly Lys Gly Phe Phe Gln Val
 85 90 95
 Thr Leu Glu Asp Lys Val His Tyr Thr Arg Ala Gly Asn Phe Arg Phe
 100 105 110
 Thr Gln Asp Gly Phe Leu Asn Asp

804260-4522000

- 53 -

115

120

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1450 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 3..414

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1083..1450

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GA TCT AAA GAG TCT ACA TAT ATT GCC CGA ATT GAA AAT TCT ACA AGT 47
Ser Lys Glu Ser Thr Tyr Ile Ala Arg Ile Glu Asn Ser Thr Ser
1 5 10 15

GAA AAA ACA CTA AAT GAT CTT GAT ATA CTT TTA AAA GAT GTG ATG TTA 95
Glu Lys Thr Leu Asn Asp Leu Asp Ile Leu Leu Lys Asp Val Met Leu
20 25 30

ACA TCA AAA AAG CAT GAA TCA CGT AGA CTT GCA GAG TCT GTA CAT CAA 143
Thr Ser Lys Lys His Glu Ser Arg Arg Leu Ala Glu Ser Val His Gln
35 40 45

AAT ATT CTT ACC CAC CTT ATA CAA AAA AAT TAT AAT ACT CAC AAT GGT 191

- 54 -

Asn Ile Leu Thr His Leu Ile Gln Lys Asn Tyr Asn Thr His Asn Gly
 50 55 60

GGG ATA AAA TCT GCA CCT TTT CAT GTT CTT ATA GGA CCC AAA ATA CCA 239
 Gly Ile Lys Ser Ala Pro Phe His Val Leu Ile Gly Pro Lys Ile Pro
 65 70 75

AGT ATT CTT GTT GAA GTA GGT TAC TGT AGT AAT AAA GCT GAA GCA CAG 287
 Ser Ile Leu Val Glu Val Gly Tyr Cys Ser Asn Lys Ala Glu Ala Gln
 80 85 90 95

CGT CTG GCA TCT AGT AAT TAC CAA AAA GCA TTA ATA GAA GGA TTA GCT 335
 Arg Leu Ala Ser Ser Asn Tyr Gln Lys Ala Leu Ile Glu Gly Leu Ala
 100 105 110

AAA GGT ATT TTC TGT TAC CTA AAA AAA CTA CAT CAC CTT GAT ATT TAC 383
 Lys Gly Ile Phe Cys Tyr Leu Lys Lys Leu His His Leu Asp Ile Tyr
 115 120 125

TCT AGT TTT ATY CTA TCT AAT TGC ACT TAA T AGCTTGGACA ATTATTATAT 434
 Ser Ser Phe Ile Leu Ser Asn Cys Thr *
 130 135

GAAGGGTATC CATGTGAAGG TACCTGGTTA AGCTTTTAAA TGTA AAAAATT ATGCAACCAT 494

ACYTTATTCC TTCAGAGGAG CTTCAATTATG AAAGTAAAAA CTCTTTCCAT GGCTATTTTA 554

GCTTGTTTAT TAGTAGCTAA CAGTGCATTT TCGGCTGACT TCCCTATTGG TGTCTTTAAT 614

TCTCAATCCA TTGCCATGGA GAGTGAAGCA GCTAAGGCCG CTCAAAAAAA ATTACAATCA 674

GAATTTGGTA ATGAAAAAAC ACAACTTGAA AACAAGCAAA AGWTTGCMMA CAAAAGCTGA 734

TGATTTACAA GCTWAGTCAG CAGCTATGTY TAACCAAGCA CGTGAAGATA AACAAAGAGA 794

ATTTCTTGAA CTTTCGTCGTA ATTTTGAAGA AAAATYTCGT GACTTTGCAA TACGTGTCGA 854

0907574.09499 864260.4252060

- 55 -

ACAAGCTGAA AACACATTAC GTCAATATNT AGCTGAACAA ATNTATNTTG CTGCTGAAAC 914

TATAGCAAAA AAGAAAGGGT TAAACTTGTT TTGATAGTGT TAGGGAAGTG TAATGTACCT 974

TGAAAAAAT TTAGATATTA CAAAGAAATT YTTGAAGCCA TAAATGCTGC ATGGAAAAA 1034

GGTGGAAGTA AACTTCCAGA GATGGCAAAC CGGAAAAAAT AACAG ATG CCC CAG TAT 1091

Met Pro Gln Tyr

1

AAA CTT TCA GAA ATT GCT AAA CTT TTA AAC TTA ACA TTA CAA GGT GAT 1139

Lys Leu Ser Glu Ile Ala Lys Leu Leu Asn Leu Thr Leu Gln Gly Asp

5 10 15 20

GAT ATT GAA GTT GTA GGC GTA AAT ACA CTT CAA GAT GCA TCA CCA AAT 1187

Asp Ile Glu Val Val Gly Val Asn Thr Leu Gln Asp Ala Ser Pro Asn

25 30 35

GAG ATA AGT TTT CTA GCA AAT GCT AAA TAT ATT CAC CAG CTT GTT TTG 1235

Glu Ile Ser Phe Leu Ala Asn Ala Lys Tyr Ile His Gln Leu Val Leu

40 45 50

TCA CAG GCT GGT GCT ATT ATT CTT TCA AAA GAA TAT GCT AGT CGT GTT 1283

Ser Gln Ala Gly Ala Ile Ile Leu Ser Lys Glu Tyr Ala Ser Arg Val

55 60 65

CCA CGA GCA CTA ATC AGT ACT GAA CCA TAT AGA GAT TTT GGT AGA GTT 1331

Pro Arg Ala Leu Ile Ser Thr Glu Pro Tyr Arg Asp Phe Gly Arg Val

70 75 80

CTT TCT TTA TTC TCT ATA CCT CAA GGA TGT TTT GAT GGT ATA AGT CAT 1379

Leu Ser Leu Phe Ser Ile Pro Gln Gly Cys Phe Asp Gly Ile Ser His

85 90 95 100

CAA GCT TAT ATA CAC CCT ACA GCA CAA GTC TCT AAA ACA GCT ACT ATC 1427

Gln Ala Tyr Ile His Pro Thr Ala Gln Val Ser Lys Thr Ala Thr Ile

105 110 115

0907574-0949864260-4252050

- 56 -

TAT CCT TTn GTT TTT ATA GGA TC

1450

Tyr Pro Xaa Val Phe Ile Gly

120

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 137 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Ser Lys Glu Ser Thr Tyr Ile Ala Arg Ile Glu Asn Ser Thr Ser Glu

1

5

10

15

Lys Thr Leu Asn Asp Leu Asp Ile Leu Leu Lys Asp Val Met Leu Thr

20

25

30

Ser Lys Lys His Glu Ser Arg Arg Leu Ala Glu Ser Val His Gln Asn

35

40

45

Ile Leu Thr His Leu Ile Gln Lys Asn Tyr Asn Thr His Asn Gly Gly

50

55

60

Ile Lys Ser Ala Pro Phe His Val Leu Ile Gly Pro Lys Ile Pro Ser

65

70

75

80

Ile Leu Val Glu Val Gly Tyr Cys Ser Asn Lys Ala Glu Ala Gln Arg

85

90

95

Leu Ala Ser Ser Asn Tyr Gln Lys Ala Leu Ile Glu Gly Leu Ala Lys

100

105

110

Gly Ile Phe Cys Tyr Leu Lys Lys Leu His His Leu Asp Ile Tyr Ser

09077574-092498

115 120 125

Ser Phe Ile Leu Ser Asn Cys Thr *

130 135

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 123 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

(xi) ~~SECRET~~

Pro Gln Tyr Lys Leu Ser Glu Ile Ala Lys Leu Leu Asn Leu Thr Leu
1 5 10 15
Gln Gly Asp Asp Ile Glu Val Val Gly Val Asn Thr Leu Gln Asp Ala
20 25 30
Ser Pro Asn Glu Ile Ser Phe Leu Ala Asn Ala Lys Tyr Ile His Gln
35 40 45
Leu Val Leu Ser Gln Ala Gly Ala Ile Ile Leu Ser Lys Glu Tyr Ala
50 55 60
Ser Arg Val Pro Arg Ala Leu Ile Ser Thr Glu Pro Tyr Arg Asp Phe
65 70 75 80
Gly Arg Val Leu Ser Leu Phe Ser Ile Pro Gln Gly Cys Phe Asp Gly
85 90 95
Ile Ser His Gln Ala Tyr Ile His Pro Thr Ala Gln Val Ser Lys Thr

- 58 -

100	105	110
Ala Thr Ile Tyr Pro	* Val Phe Ile Gly	
115	120	

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 559 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..557

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GA TCA AAG CCG CAT TTA CNG CAA GAG TCA GAA ATT GAA GTT TTG AAA	47
Ser Lys Pro His Leu Xaa Gln Glu Leu Glu Ile Glu Val Leu Lys	
1 5 10 15	
AAA GAA GAC TTT GGG CGT CAT ATT GTT AAA TTA TGC TGG AAA GGT TCT	95
Lys Glu Asp Phe Gly Arg His Ile Val Lys Leu Cys Trp Lys Gly Ser	
20 25 30	
TTA TCA AAT ATC TTT TTT TCC TAT GGG GAT ATC CCG CAC CCA CCT TAT	143
Leu Ser Asn Ile Phe Phe Ser Tyr Gly Asp Ile Pro His Pro Pro Tyr	
35 40 45	
ATA CAT CAA AGT AAT AAG GTT CAG GAT AAG GAA AGA TAT CNT ACN GTA	191

000754 09249 864260 4252060

WO 97/20050

- 59 -

Ile His Gln Ser Asn Lys Val Gln Asp Lys Glu Arg Tyr Xaa Xaa Val
 50 55 60

TAC TCT ATA TTA CAT AAN CTG GGT TCT GTA GCA GCT CCT ACA GCT GGA 239
 Tyr Ser Ile Leu His Xaa Leu Gly Ser Val Ala Ala Pro Thr Ala Gly
 65 70 75

TTA CNC TTT TCT GAA ACT AGC CGT NAT AAA TTA CAC AAA NAT GGT ATT 287
 Leu Xaa Phe Ser Glu Thr Ser Arg Xaa Lys Leu His Lys Xaa Gly Ile
 80 85 90 95

AGT TGG GCA TAA ATC CCT CTT CAC GTG GGA TAT GGA ACA TTC AGT CCC 335
 Ser Trp Ala * Ile Pro Leu His Val Gly Tyr Gly Thr Phe Ser Pro
 100 105 110

GTC CTC TGC AAT GAC ATC CCA AAA CAT CTT ATC CNT TCT GAG TTT GTT 383
 Val Leu Cys Asn Asp Ile Pro Lys His Leu Ile Xaa Ser Glu Phe Val
 115 120 125

CAC TTT CCT GAA ACT ACN TTT TCC ACT ATA TTA AAT GCA CGG TTT GCA 431
 His Phe Pro Glu Thr Xaa Phe Ser Thr Ile Leu Asn Ala Arg Phe Ala
 130 135 140

NGG GAA TAC CTA TGT TCT GCC ATA GGG GAC CCA CTG TTG TCC CCA CCA 479
 Xaa Glu Tyr Leu Cys Ser Ala Ile Gly Asp Pro Leu Leu Ser Pro Pro
 145 150 155

TTG GAN GGG TGT TAT CTT ACC CCT TTC GCC CGG GGT TCC CCT CCC CAA 527
 Leu Xaa Gly Cys Tyr Leu Thr Pro Phe Ala Arg Gly Ser Pro Pro Gln
 160 165 170 175

CCC TAT TCC ATT GNG TTT TCC TCT CAA ATT AT 559
 Pro Tyr Ser Ile Xaa Phe Ser Ser Gln Ile
 180 185

00077574-00496

- 60 -

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 185 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Lys Pro His Leu Xaa Gln Glu Leu Glu Ile Glu Val Leu Lys Lys
 1 5 10 15
 Glu Asp Phe Gly Arg His Ile Val Lys Leu Cys Trp Lys Gly Ser Leu
 20 25 30
 Ser Asn Ile Phe Phe Ser Tyr Gly Asp Ile Pro His Pro Pro Tyr Ile
 35 40 45
 His Gln Ser Asn Lys Val Gln Asp Lys Glu Arg Tyr Xaa Xaa Val Tyr
 50 55 60
 Ser Ile Leu His Xaa Leu Gly Ser Val Ala Ala Pro Thr Ala Gly Leu
 65 70 75 80
 Xaa Phe Ser Glu Thr Ser Arg Xaa Lys Leu His Lys Xaa Gly Ile Ser
 85 90 95
 Trp Ala * Ile Pro Leu His Val Gly Tyr Gly Thr Phe Ser Pro Val
 100 105 110
 Leu Cys Asn Asp Ile Pro Lys His Leu Ile Xaa Ser Glu Phe Val His
 115 120 125
 Phe Pro Glu Thr Xaa Phe Ser Thr Ile Leu Asn Ala Arg Phe Ala Xaa
 130 135 140

09077574-092498

WO 97/20050

- 61 -

Glu Tyr Leu Cys Ser Ala Ile Gly Asp Pro Leu L u Ser Pro Pro Leu
 145 150 155 160

Xaa Gly Cys Tyr Leu Thr Pro Phe Ala Arg Gly Ser Pro Pro Gln Pro
 165 170 175

Tyr Ser Ile Xaa Phe Ser Ser Gln Ile
 180 185

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 477 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..294

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

T ATA AAA CAT TAG CGN CTT TNG TAT TTG GAC TTC AAA AAA ATT TTT
 Ile Lys His * * Leu * Tyr Leu Asp Phe Lys Lys Ile Phe
 1 5 10 15

AAT TAT ATA GGA GAA CAT TCA CCA TTA AAA CGT AAT GTA ANT ATG GAA
 Asn Tyr Ile Gly Glu His Ser Pro Leu Lys Arg Asn Val * Met Glu
 20 25 30

GAT GTA GGT AAA TCT GCT GTT TTT TTA GCT TCA GAC CTN TCA TCA GGA
 Asp Val Gly Lys Ser Ala Val Phe Leu Ala Ser Asp * Ser Ser Gly
 142

46

94

- 62 -

35	40	45	
GTA ACC GGT GAA TTN TTT TTG TTG ATG CTG GNA CAA TAA TTT AGG TAT			190
Val Thr Gly Glu * Phe Leu Leu Met Leu * Gln * Phe Arg Tyr			
50	55	60	
TTA ACC ATA CAT GCT TTA TAC AAC ATA TTG TGA GTT ACA ATA GCC ATA			238
Leu Thr Ile His Ala Leu Tyr Asn Ile Leu * Val Thr Ile Ala Ile			
65	70	75	
ACA CAT TTA TAT TCT ATA TAA TAA CAG TAG AAT AAT AAT AGA ATA TTT			286
Thr His Leu Tyr Ser Ile * * Gln * Asn Asn Asn Arg Ile Phe			
80	85	90	95
TTT ATG ACC ATTTGTATCT ATACAATAGT AAATAGATTA ATACATATAA GACTATATTC			344
Phe Met Thr			
TTTTTGAGAG CAACTTAAAG GAGCGGTTAT GGCTTTAGTT ACAAAAAGAAG AAGTACTTCA			404
ATACCATAGT GAACCCCGAC CAGGTAAACT TGAAGTATTT TCTATAAAAC CATGTAAAC			464
ACAAAAAGAT CC			477

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ile Lys His * Xaa Leu Xaa Tyr Leu Asp Ph Lys Lys Ile Phe Asn

(2) INFORMATION FOR SEQ ID NO:15:

(ii) MOLECULE TYPE: DNA

- (ix) FEATURE:

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

WO 97/20050

- 64 -

G GAA TTG TTA GTA TTC TCC CAG AAC AGA AGC CAA AAT ATT TGG CTA 46
 Glu Leu Leu Val Phe Ser Gln Asn Arg Ser Gln Asn Ile Trp Leu 15
 1 5 10

CTT ACA TTA CCT ATT TTT GTG TTA GGT ATA GCA CAA GGT ATA TCA TTT 94
 Leu Thr Leu Pro Ile Phe Val Leu Gly Ile Ala Gln Gly Ile Ser Phe 30
 20 25

CCT TTA GTA AAC AGC CAC ATT ACA TCA CTT GCA CCA ACA TCC AAC AGA 140
 Pro Leu Val Asn Ser His Ile Thr Ser Leu Ala Pro Thr Ser Asn Arg 45
 35 40

GCT ATT GTT ATG GCT ATA AAC AGT ACA TTT ATG AGG TTA AGT CAG AGT 190
 Ala Ile Val Met Ala Ile Asn Ser Thr Phe Met Arg Leu Ser Gln Ser 60
 50 55

ATT TCG CAA ATG GTT TTT GGT ATT GGA TGG TCA TTT TTT GGT TGG CCT 238
 Ile Ser Gln Met Val Phe Gly Ile Gly Trp Ser Phe Phe Gly Trp Pro 75
 65 70

GGT CCT TTT ATA TTT GGT CTT TTT ACT TCT ATT ATA TTA GCC CTC TTA 286
 Gly Pro Phe Ile Phe Gly Leu Phe Thr Ser Ile Ile Leu Ala Leu Leu 95
 80 85 90

ATT ATG AAG TAT TTT CAA GAT GTA ACC CAA TAT CAC CTA TTT TTG ATA 334
 Ile Met Lys Tyr Phe Gln Asp Val Thr Gln Tyr His Leu Phe Leu Ile 110
 100 105

AGT AGT AAA TTT TAT TAT TAA AAA GCT TAG TTA GTT AAG ATT ACA TAT 382
 Ser Ser Lys Phe Tyr Tyr * Lys Ala * Leu Val Lys Ile Thr Tyr 125
 115 120

ATT ATA TAC AAT TAC TAT AAC ATT AAC TAA TTA CTA ACT ATT ACT TCC 430
 Ile Ile Tyr Asn Tyr Tyr Asn Ile Asn * Leu Leu Thr Ile Thr Ser 140
 130 135

AAT TGA TTA ATT GAT GCT ATT TAA AGA GGA TAT ATT AAT GAT GTC ATG 478

007754-092499

- 65 -

(2) INFORMATION FOR SEQ ID NO:16:

(A) LENGTH: 174 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

(x1) SEQUENCE

Glu Leu Leu Val Phe Ser Gln Asn Arg Ser Gln Asn Ile Trp Leu Leu

1 5 10 15

1
Thr Leu Pro Ile Phe Val Leu Gly Ile Ala Gln Gly Ile Ser Phe Pro
20 25 30

Leu Val Asn Ser His Ile Thr Ser Leu Ala Pro Thr Ser Asn Arg Ala
35 40 45

Ile Val Met Ala Ile Asn Ser Thr Phe Met Arg Leu Ser Gln Ser Ile

50
Ser Gln Met Val Phe Gly Ile Gly Trp Ser Phe Phe Gly Trp Pro Gly
65 70 75 80

65
Pro Phe Ile Phe Gly Leu Phe Thr Ser Ile Il Leu Ala Leu Leu Ile
85 90 95

- 66 -

Met Lys Tyr Phe Gln Asp Val Thr Gln Tyr His Leu Phe Leu Ile Ser
 100 105 110

Ser Lys Phe Tyr Tyr * Lys Ala * Leu Val Lys Ile Thr Tyr Ile
 115 120 125

Ile Tyr Asn Tyr Tyr Asn Ile Asn * Leu Leu Thr Ile Thr Ser Asn
 130 135 140

* Leu Ile Asp Ala Ile * Arg Gly Tyr Ile Asn Asp Val Met Ala
 145 150 155 160

His Asn Arg Cys Tyr Pro Trp Ile Ser Ala Trp Asp Pro Gly
 165 170

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 846 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TATTTACTCG CGCGGCCGGG CGTCTTACAC AAATGGATCC CTTGCANTAA TCCAAGGATA	60
ACNCCTATTG TGANCCATGA ACATCATCAN NATATCCTCT TTANATAGCA TCNANNNNNTC	120
AANNGGAATT AACAGTTACT ANNTAGTTAA TGTCATAGTA ATTGTCNATA ATATATGTAA	180
TCTTAACTAA CTAAGCTNNT TAATAATAAA ATTNACTACT TATCAANAAT AGGTGATATN	240
GGGTTACATC TTGAAAATAC TTNCCATAAT TANGAGGGCT AATATAATNG AANTAAAAAG	300
ACCANATATA AAAGGACCAG GCCAACCAAA AAATGACCAT CCAATACCNA AAACAATTGG	360

09077574-092498

- 67 -

CGAAAATACT CTGACTTAAC CTCANAAATG TACTGTTTAT AGCCATATCA ATAGCTCTGT 420
 TGGATGTNGG NGCAATTGAT GTAATGTGGC TGTNTACTAN ANGAAATGAT NTACCTCGTG 480
 CTATNCCTAN NACAANAATA NGTAATGTAA GTANCCNAAT ATCTTGGCTT TGTAATGGGA 540
 GAATAATNNC AAGTCCTTGG GAAATNAANT TACNNCCAGC CAGCTATNNT AAGCAGTTCT 600
 NTGGTGACTA TACGTCCTAC TNAANTCGTG CCAAAGATTA AATANNCGAT AATCGCNCNTN 660
 CCTAAANCAN GCAATACTAA AATGGTTTCT NCCTANCTTG GNATANGGTG GAAGCNCGGA 720
 CAGAATTNAN TTCGCNANTT TANANNGGAA NATNCGTNAA NTTANTCGGG GCCCANNCCN 780
 AAATTCCTNA NTCNATANAN NAACTNNCTN CTNTAAAANG GCCNACTGGA NTNGTTAAAT 840
 GAAATA 846

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 855 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GATTNTTTAT CGATCACTNT AGACGCGATT TGGGNAACAC TTACCTGGTA NCCACCCGGG 60
 TGGAAAAATC GATGGGCCCG CGGCCGCTCT AGAAGTACTC TCGAGAAGCT TTTTGAATTC 120
 TTTGGATCCT CAACACAGGG TATGGATTAA AACAACCTTA GCTCTAACAG GAGCATTTTA 180
 TAATATATTC CCTGGTAGAA CAATATCTAC TCAAGAAAAT CTGTCTATTG GTTTTCAACT 240
 AAAAAAACT TTAAACCTT TTCATTGGAC CATCTTACTC TTAGATGAAC ATTATATGTC 300
 TTCGCAAGA ATTGCAGCAG CAATTATGCC TGCACAGCTT GCTGGAGTTA AAAACATTAT 360

864260-4254060

- 68 -

AGCTGTTTGG ACCAGTAAAA ATAACCGACT GACCGCTGAA AAAATCTCAC CTGCTTTACT	420
AACAACATTA GAACTTTCAG GAGTTAACAT AGCCCTAACA CTTACCCACA CTGAAACTGA	480
ACTTCTTATT CATCAATTAA TGAAAATAGG TATTGGAAAC CTGTTATATT TTTTAAAAGA	540
AGAAGACATA CTACATATAT CTACTATACC TGTACTACCT TTCTGGAAAG AATATACTTC	600
TCATCGACTT GTTATAGAAA AAGATGCTGG CNTTAATACA GAAATCCTCC AATGGGCNCA	660
TCCTCATTCA ATTATTGAAC AAATAGCAAC AGAACCATAC TCTGAAANAT ATCCCAGATG	720
CACTTTACTG TGCTAGCTCA TCCANTAAAA ACTATNCTCA TANAGNATCC CCAGAATTTT	780
TCATNATGGA CTGGAACCTA TTTGGATTCA NCCCAACNCT TCCTCCAANC CTCCTTTCTC	840
CATACACCAT GGGGA	855

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1082 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TATCTNGTTG ANTCAATAAA ACTTTTGGGG CCCNTNAAAN TTTCATNANN AAAAAACAA	60
NATTNCTGGG GGNCCCNTCC CAAAAAANNC AATCANTNNG AANCTTGNCT TCTTATTNNG	120
NTTTTNANAC TATAATATNT NTTATCNATA ATNNATCNNT ATACTNATTT CTNATTCANT	180
NACANNGGNN AGNAANNTTA ATCTNAAANA CTNCNAAGGG GGNNNTNATA NTNTTTNTTT	240
NTTTNTCCCN TNNAATNNAT AACCNNNCAC CCNNATTANT TNNAATNNAT ACCATANCNN	300
CCTTTCAAAC TGTACACATA NTANNNAANN ACACTCNANC NTTTTNCATC CTCTCTANTN	360

09077574-092498

- 69 -

CCNACTCCNA TNNANCTNTT CCCCCATNCC TATNTINTCNC TGCTTCCCAG NTTNNACNTN 420
 NCTTNNTTTC ACANTATTCC TATCCAANCT AACATNTNTN NTNTCNTNCT CCTTNTNTNT 480
 TATNTNTTTC TNNTACCTNN CACTGACANT CTATNANTNA NNTCNNATAC TNNTATANCT 540
 NTANGCNANT NTATCTANAA NTNTANCNNN NNATCNTNAC NGCCGTNNAT NTNNNNNCAN 600
 TTANNTANNN CTANCNTNNC CAANNNCNTA TNTATNAATA ACNACTATCC NATATTNNAT 660
 TMNTNTNTNT CNTANNCAAA TNATTTANGC NCACNNCACT ANGNTATATN ANNATTNTAT 720
 ATTNTGAANC TTCTNGGCTT CNCNAATANT ACCANTNNNC ANCNTCNMNT NCATCTNNNT 780
 NTACTTCNTA CCATANCGCT CTCNAGNNTC ACTACTTCTA NTAGTNATCH TCTACTGCCN 840
 ATGGCNNNNN GCNNNNCGAN AGNTATNCAC NTACANTNNC NTCTACTATN TANATCTANN 900
 NCNTCCGNNG CCTNCNGTAC GNNTNGGCNA ANTCGNNTAC TTTNCNTNTA TCTAGTCNCA 960
 TCAGNNNTNG ANTCCTCAAN CNNGCTCTAN TTACATGTNN MNTNATGCNC TANANCGNNA 1020
 CNTCTATCCT TCNANTCTGC NCTNANTNTA TANACTCTNN MNNATCNNCN AANCTATNTC 1080
 CC 1082

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 354 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:20

CTCCCNINNC NCTAAGTGGA NTCGCGCGCT GCAGGTCGAC ACTAGTGGAT CTTGATATAC 60
 TTTTAAAAGA TGTGATGTTA ACATCAAAAA AGCATGAATC ACGTTAGACT TGCAGAGTCT 120
 GTACATCAAA ATATTCTTTA CCCACCTTAA TACGAAAANA AATNNTTATN CNCCNCNATG 180
 GGTGGGGNTN AAATCCTNGC CCCNTTNCCC TGTTCTTTTA GGGAACCCCC NAATTCCCCN 240
 NGTTATTCTT CTGTTTGAAA NTTCTGGTTN CCCGGCCCTN TNACCAANAG CTTGANNNCC 300
 NCCCCGTCCT GGGGCATCCT CNTGTTTATT TTCCCTCNAN CNCCCCCTTN ACTN 354

090754-09498

- 70 -

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 477 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:21

GGATCTTTTT GTGTTTTACA TGGTTTTATA GGAAATACTT CAAGTTTACC TGGTCGGGGT	60
TCACTATGGT ATTGAAGTAC TTCTTCTTTT GTNACTAAAG CCATAACCGC TCCTTTAAGT	120
TGTTCTCAAA AAGAATATAG TCTTATATGT ATTAATCTAT TTACTATTGT ATAGATACAA	180
TAGGTCATAA AAAATATTCT ATTATTATTC TACTGTTATT ATATAGAATA TAAATGTGTT	240
ATGGCTATTG TAACTCACAA TATGTTGTAT AAAGCATGTA TGGTTAAATA CCTAAATTAT	300
TGTNCCAGCA TCAACAAAAA NAATTCACCG GTTACTCCTG ATGANAGGTC TGAAGCTAAA	360
AAAACAGCAG ATTTACCTAC ATCTTCATA NTTACATTAC GTTTTAATGG TGAATGTTCT	420
CCTATATAAT TAAAAATTTT TTTGAAGTCC AAATACNAAA GNCGCTAATG TTTTATA	477

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 568 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:22

09077574-09499
964260-42522060

- 71 -

GATCATTTAA	AAAACCATCT	TGAGTAAAAC	GAAAATTCCC	TGCTCGTGTA	TAGTGACTT	60
TATCCTCTAA	TGTAACCTGA	AAAAAACCTT	TTCCACCAAT	AGCAAGATCT	GTTACACTAT	120
TGCCAGGTTC	AAAAGCACCC	TGTGTAAAAA	TTGTGCGAAC	ACTTCCAACC	TGTGCTCCCA	180
TACCAGCCTG	GTTTGGCCCC	TGACTTCCAG	TAAAACCTAT	TGCTAAATCT	TGACTAAACA	240
GGTCTTGAAA	CACTACCTGT	TGCTGCTTAT	ACCCAATGGT	ATTTGCGTTA	GCAATATTAT	300
TGGAGACAGT	ACCANCCCTG	TNCTATGGGT	TTTCATACCT	GTTGGCANCA	ATAAACAAAC	360
TCCCCATCAT	GATAACATCT	CCTAAAAAAT	AATTTTCATGG	NGGNAAAAAT	GTTACCTACA	420
CATCTCTATT	TTNAAAGCAA	AAAACCCATG	CCCAANAAAA	TTTTTGCGCC	NAATTAATAT	480
ACTTAATCTA	ATAAACTTTT	TTGGGTAATN	AAAAAAAATT	AATTTTTTTAA	ACTTGGTTCN	540
ACCAACCTTT	TCTCCTTACT	TTTTAACC				

09077574.092498

- 72 -

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 477 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:23

GGTACCCAC CCGCTGGAA AATCGATGGG CCGCGGCCG CTCTAAAANT	50
ACTCTCGAGA AGCTTTTGA ATTCTTTGGA TCCCCAGGAA TAACTTGTTG	100
ACGGAATTTT ACATTTTCTA TCCCTGCAA TANAAAAACT TTACCTTGTA	150
G TTCATTAAT AGGAAAAGAT TGGAGTACTG TGATTCCACC TGATTGCGCC	200
ATAGCTTCTA AAATTAGAAC TCCAGGCATG ACAGGAAATC CAGGGGAAAT	250
GACCCNGAAA AAATGGTTCA TTAATACTAA CATTTTATA AGCTTTAATA	300
TATTTGCCAG CATTAAATTC AATAACTCTA TCTACAATTA AAAAGGGATA	350
ACGGTGGGGA ATTTACTGTA AAATTTCTTG GATATTTTGG AGGTATGGAT	400
GGGGACATTA ATTTTCCTAT ATATATGCTC TTTTCTTTT CNAAAATTTT	450
TCAGCTTTTT TATCCCNTAA AAACCTC	467

09077574.092499
864260.4252060